

App No.: NYA
Docket No.: HO-P03371US0
Inventor: Sandra E. Lavens et al.
Title: 2031 OXIDOREDUCTASE

	201	211	221	231	241	251	261	271	281	291
	-----6-----	-----	-----	-----	-----7-----	-----	-----	-----	-----8-----	-----
SEQ 3	LAHAGRKATT	VAPW	-----ISPS	AIATEKVGOG	PDRVKGPGDI	P	-----FABFPAPK	KA	-----MTLDBIE	QPKK-DWAA
SEQ 6	IAHAGRKAST	VAPW	-----LSAN	DTAEKMGCH	PDIVKGPSTV	P	-----FTDGNVFP	KE	-----MTQDIE	DLKT-AMVA
SEQ 8	LAHAGRKASA	VAPW	-----LAAQAQKSS	LKADESVOGW	PADVVGPSGG	E	-----BHIP	SPBEDAYW	RA	-----LSTAEVR
SEQ 10	IAHAGRKASIN	IAPW	-----LMMKG	IVATEKVGOG	PDVIGPSTV	P	-----FHEFTPTP	KA	-----MTKDDIE	QPKR-DWPA
SEQ 12	LGHGGRKASG	QPLF	-----LHLE	QVADKSVNGF	ADKAVAPSAL	A	-----FRPHGNLF	VPNE	-----LTQDEIK	RVVK-DFGA
SEQ 14	LAHAGRKAST	LAPW	-----QIQHGW	QEHCVGPSTE	P	-----PSDSHPTP	RE	-----LTNENIN	SIVE-DFANA	
SEQ 16	LAHAGRKAST	KAPW	-----HYQRGKS	ELAGPEQGGW	PENVWAPSAL	S	-----YNEETFPFP	KE	-----MTVEQIH	SLVE-AMKA
SEQ 19	INHPRQSPM	GAGT	-----RGW	E-KAVAPSVP	P	-----LVLGFAFVP	RLLSKVLPGT	PRELTVARIK	DIV-QKPAVT	
SEQ 22	LAHAGRKAST	KAPWHDSTPT	SGEYKPREGL	QVVGPEYGGW	PDDVWAPSAL	P	-----FSEDFFNP	KE	-----MTVEBIE	GLVT-SFVDA
SEQ 24	LAHAGRKAST	TAPY	-----RG-Y	TVATEAQGGW	ENDVVGPPFN	E	-----DRMDENHAQP	HK	-----LTERQYD	SLVD-KFVA
SEQ 27	LAHAGRKAST	LAPW	-----ITEARGK	ALAQESSEGW	PDDVWAPSAL	P	-----VTKDWAFT	RE	-----LTTE	SR
SEQ 30	LSHAGRKASC	VSPW	-----LSVN	AVAAEEVGGW	PDNIWAPSAL	A	-----QENGVNFPV	KA	-----PTKEDIE	QLKS-DYVEA
SEQ 33										
SEQ 35	IGHAGRKAST	VVPW	-----LDRK	NTAF?						
SEQ 38	LSHAGRKASC	VSPW	-----LSIN	AVAAKEVGGW	PDNIWAPSAL	A	-----QEAGVNFVP	KA	-----PTKEDIE	ELQN-DFLAA
SEQ 40	LAHAGRKAST	VAPW	-----LSGG	DVAGEDVNGW	PDDVWAPSAL	P	-----WNEKHAVP	KE	-----MSLDDIE	APKK-APGEA
SEQ 42										
SEQ 44	VGHPRQARG	SVQ	-----QHPISASD	VOLKQEM			-----FGSKFGVP	RP	-----ATKEDIK	AVIE-GFAHT
SEQ 83	IGHAGRKASC	VAPW	-----LDAG	LAEXKAAGGW	PDDVVGPSNE	P	-----FAPGYPTP	RA	-----ITLEBIE	QLKE-DFVSG
SEQ 85	LAHAGRKASD	WSPF	-----YRGEKKQ	KFTVQDEGGW	PDRVWAPSAL	A	-----YAQGHVFT	RA	-----LTTEIDIN	KLQD-KFVQS
Bacteria										
T4612	IAHAGRKASA	NRPW	-----EGDD	HIGADDARGW	ETIAPSAI	A	-----PGAHLFNV	PRA	-----MTLDDIA	RVKQ-DFVDA
NP_625402	LAHAGRKAST	AQPW	-----RGG	APVGADAYGW	OPLAPSAI	A	-----FDERHFPV	TE	-----LTVPQIQ	EAVG-RFADA
NP_295913	LAHAGRKAST	YAPW	-----RGK	GAVPAELGGW	QVIGPDEN	S	-----PHDLFPTP	AM	-----MGADBLR	GVVD-APSAA
AF320254	LGHAGRKGAT	KLAW	-----EG	IDEPLEAGAW	ELISASPL	P	-----YLPHSQVP	RA	-----MTRDDME	RVRN-DFVRA
OYE family										
AF4875	LMATGRAADP	DVLA	-----DMK-D	LISSS-AVPV	SEKGP		-----LP	RA	-----LTEDBIO	QCTA-DFQAQ
AF4961	LMHARATIP	QMTG	-----SPAVAS	ATVMSPTTC	YSHIP		-----VGST	EPVRYADHPP	IE	-----LTIP-HL
C62460	LMYLGRVANP	KDLK	-----DAGLPL	IGPSA-VYM	DEESE			KLAKSVGNEL	RE	-----LTKEBID
Nc4452	LMHSLGRAANP	DVLA	-----KBGLK	LKSSS-AVPV	EXEAP			VP	EE	-----MTVAEIK
ScOYE1	LMVLGWAAPF	DNLA	-----RDG-LR	YDSASDNVYM	DAEQS			AKAKKANNPO	HS	-----LTQDEIK
ScOYE2	LMVLGWAAPF	DTLA	-----RDG-LR	YDSASDNVYM	NAEQS			EKAKKANNPO	HS	-----ITKDEIK
ScOYE3	LMVLGWAAPF	DVLA	-----RDG-LR	YDCASDRVYM	NATLQ			EKAKDANNLE	HS	-----LTQDDIK
A36990	LMYLGRVANA	KDLK	-----DSG-LP	LIAPS-AVYM	DENSE			KLAKKAGNEL	RA	-----LTEDBIO
SEQ 3	TKRAIAA-GA	DFVEIHNAHG	YLLSSFLSP	AANNRTDOY	G-GSFENRIR	LSLEIAQLTR	DAVGPHVP	VPLR	ISAS-DWCE	ETLPEQ
SEQ 6	VGRAVKA-GA	DFVEIHNAHG	YLLMSFLSP	AVNTRTDEY	G-GSFENRIR	LSLEIAKLTR	ENVPKDMP	VPLR	VSAT-DWLE	EVOPNKP
SEQ 8	ARLAVQA-GV	DFVEIHNAHG	YLINEFLSP	VTNKRTRDAY	G-GSFENRIR	IVREVAIAIR	AVIPEGMP	LPLR	ISAT-DWLE	GQPVAAESG
SEQ 10	CKRAIAA-GA	DFVEIHNAHG	YLLSSFLSP	SNTRTDEY	G-GSFENRIR	LSLEIAQVTR	DAVGPNVP	VPLR	VSAT-DWIE	ETLPEE
SEQ 12	ARRAVEISGF	DAVEIHGAHG	YLINEFPYSP	ISNKRTRDEY	G-GSFENRIR	FLKEVIDSVK	SSIPNDVP	VPLR	ISAA-ENSP	DPE
SEQ 14	AMRAVEISKF	DAIEIHGAHG	CLIHQFLSK	LTNKRADQY	G-GSFENRIR	FLAQIENIK	RKIET	IFLK	FPMS-DNCS	DPE
SEQ 16	AQRALKA-GF	DLIEIHNAHG	YLISEFLSP	ISNKRTRDOY	G-GSFENRIR	VLRISIAVR	SVIPEMP	LPVR	VSAT-EMME	YTGGP
SEQ 19	ARITADA-GF	NGVEIHNAHG	YLLAQFLSK	KTNRGDEY	G-GSLENRR	IVGSIKECR	ROVTEAGVE	BAKPKVGIK	LSAA-DWAQ	GRDKKEE
SEQ 22	AKRAIDA-GV	DIYEHGAHG	YLIETFLSP	LSNKRTRDKY	G-GSFENRIR	VLIDIIKAVR	AVIPEEM	PLPVR	ISAT-EMME	YAGEP
SEQ 24	AKRAVEI-GF	DFVEIHGAHG	YLISSVSPA	PTTNDNRDKY	G-GTFEKRL	FPMEVVHSVR	KAIPDSMP	LPYR	VTAT-DWLP	KGO
SEQ 27	AKRSNRA-GF	DFVEIHNAHG								
SEQ 30	AKRAIHA-GF	DFVEIHNAHG	YLLHQFLSP	VSNQRTDEY	G-GSFENRIR	VVLEILDILIR	AAIPETTP	VLVR	VSAT-DWFP	DSQFKDEPPE
SEQ 33										
SEQ 35										
SEQ 38	KRA-RA-GF	DFVEIHNAHG	Y-LHQFLSP	VSNQRTDEY	G-GSFENRIR	VVLEII				
SEQ 40	VGRALKA-GF	DFVEIHNAHG	YLLHEPCL	RATPGTST	G-GSWENRTR	LTMSKRRCPC	QHP	LPLR	LSST-EMME	DTDQKPKP
SEQ 42	AKRAVKA-GV	DFVEIHGAHG	YLLHEFLSP	ITNKRTRDSY	G-GSFENRIR	LLIEIVIAVR	AAMPSSMP	ILGK	INSV-EPOE	KG
SEQ 44	AEYLKRA-GF	DGVEIHLAAGH	YLLAQFLSE	TTNKRTRDEY	G-GSLENRR	LILEVTAERV	RRTSKNF	LPTR	ISGT-DWLE	NNPEYBGE
SEQ 83	VRAVEA-GF	DTIDPFAHG	YLVSSFLSP	ATNKRTRDKY	G-GSLENRR	LALIEVIAAR	AVMPEDMP	LMVR	VST-DWAD	QAHQAD
SEQ 85	ARNAFEA-GY	DYVELHSAHG	YLMVHFLSP	LTNQRTRDEY	G-GSLENRR	FLINVAARRIR	QEPNKG			
Bacteria										
T4612	ARRARDA-GF	EWELHFAHG	YLGQSFFSE	HSNKRTRDAY	G-GSPDNRSR	FLLETLAAVR	EVMPENLP	LTAR	FGVL-EYDG	RD
NP_625402	ARRALAA-GF	BIAEIHGAHG	YLIHEFLSP	HSNKRTRDAY	G-GSYANRTR	FALEVVDVAVR	EVMPDDKP	LPFR	VSAT-DWLE	EG
NP_295913	ARRAQVA-GF	DAVEIHNAHG	YLLHQFLSP	LAMTRTDDY	G-GSFENRIR	LILEVVRVAVR	HYWPAHL	LPVR	LSAT-DWAE	G
AF320254	TRMAAFA-GF	DILELHCAHG	YLLSSFLSP	LTNKRTRDEF	G-GDLENRR	FPLEVFKAMR	AMMPTNRP	MSVR	LSCH-DWFP	G
OYE family										
AF4875	ARNAINA-GP	DGVEIHGAHG	YLIQPTOK	SCNHRQDRM	G-GSIENRR	FAVEVTRAVI	EAVGADR	VGK	LSPY-SQYL	GMGTMD
AF4961	AKTAMEI-GF	DGVEIHGAHG	YLPDPLSS	VNKRTRDEY	G-GSPKRCR	FVLEIMDELA	ATVGEDN	LAIR	LSPP-GLFN	QARG
C62460	AKRAISA-GF	DYVEIHSAPG	YFLDQFLNP	ASNKRTRDKY	G-GSIENRR	LLLRIDKLI	GIVGAEK	LAIR	LAPW-SSFL	GMEIEG
Nc4452	AKNAVEA-GF	DGVEIHGAHG	YLIQFLQD	TCNQRTRDEY	G-GSIENRR	FAHEVVKAVR	EAVGAEK	TGIR	LSPY-STFP	GMKMKK
ScOYE1	AKNSIAA-GA	DGVEIHSANG	YLLNQFLDP	HSNKRTRDEY	G-GSIENRR	FTLEVVDALV	EAIGHK	VGIR	LSPY-GVFN	SMSGGA
ScOYE2	AKNSIAA-GA	DGVEIHSANG	YLLNQFLDP	HSNKRTRDEY	G-GSIENRR	FTLEVVDVAV	DAIGPEK	VGIR	LSPY-GVFN	SMSGGA
ScOYE3	AKNSIAA-GA	DGVEIHSANG	YLLNQFLDP	HSNKRTRDEY	G-GSIENRR	FTLEVVDALI	ETIGPER	VGIR	LSPY-GTFN	SMSGGA
A36990	AKNALEA-GF	DYVEIHSANG	YLLDQFLNL	ASNKRTRDKY	G-GSIENRR	LLLRVVDKLI	EVVGANR	LALR	LSPW-ASFP	GMEIEG

	401	411	421	431	441	451	461	471	481	491
	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 3	SWKSEDTVR	PAQELVVK-Q	GAVDLIDISS	GGVLAQO	-----	KSGPAPQVPP	AVAVKKAAGD	-----	KLLVAAV	GAIT
SEQ 6	SWRGVDTVR	FAKILA-BT	GYVDVLDVSS	GGTHSEQ	-----	HAKPGFOAPP	AIAYKNAVGD	-----	KLAVASV	GMAI
SEQ 10	SWKMD-QSSL	ELVKRLP-E	WGIDLLDVSS	AGNHKDO	-----	NLHPTAQTEL	AQQRQAI	-----	RAAGAST	LGVAVGLITD
SEQ 12	SWKLEDSVR	FAEALAA-Q	GAIDLLDVSS	GGVHAAQ	-----	KSGPAPQAPP	AVAIKKAAGD	-----	KLLVATV	GTIT
SEQ 14	AWTIEDSKK	LADILV-E	KGIALVDVSS	GGNDYROPP	-----	RSGISK	ELREPIHVPL	-----	KLLVSCV	GGLB
SEQ 16	AWSTEDALK	LADLVI-D	LGVKVIDVTS	GGNVAHCKS	-----	RYLLAND	DKQLPSOVPL	-----	RCLLIACS	GGLD
SEQ 14	SWDLQQT	ELAKILP-D	LGVDLLDVSS	GGNNKDO	-----	KI	NVHTYQIDM	-----	AGKQLLVGAV	GLVT
SEQ 19	TDTAEVLK	QIELFE-Q	WGIDFVEVSG	GSYEDPQMAN	-----	GPKPEKSERT	MAREAFFLEF	-----	AKIIRTK	GGFR
SEQ 22	SWDLEQSTQ	LAKLLP-D	LGVDLLDVSS	GGNSVAQ	-----	KI	ELTPYYQIDL	-----	AAKIREAVGD	RLIGAV
SEQ 24	GWEIEDTVAP	TLAARLR-D	GGVDLLDVSS	GGNHKDO	-----	RI	EVKDCYQVPP	-----	AEKIKDQVNG	ILIGAV
SEQ 27	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 30	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 33	SWTVEGTC	QLARILP-K	HGVDLLDVSS	GQIHPSK	-----	AIAI	KSGPAYQVDL	-----	AKQVKAAGD	SVLVSAY
SEQ 35	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 38	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 40	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 42	SWDVESTIK	ISKILA-D	LGVDLLDVSS	GGNHPOQ	-----	KI	NMFNT	-----	-----	-----
SEQ 44	FKP-EBAVQ	LCEALEAAGM	DPVETSG	GTYESFG	-----	FAHRKESS	RKRENYPIEF	-----	ARVIRKAVKH	MVYTTG
SEQ 83	TWTLQESIK	LAHOLA-D	RGVDVLDVSS	GQIHKMO	-----	KV	AAGPGYQAPL	-----	AKAIKKSVDG	KMLISTV
SEQ 85	SWTVDTQVTE	LAKMQLB	ARVDLLDVSS	GGLVPPQ	-----	KI	TVGAGYQLFG	-----	AKAVRDALAK	IBPDASKR
Bacteria	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
T44612	EQTLESIE	ELARRFK-A	GGDLDSVSV	GFTIPET	-----	NI	PMGPAKMGPI	-----	AERVREAKL	PVTSAM
NP_625402	GWTPTDQVTR	FARDLR-A	HGIDLLDVST	GGNVPRV	-----	RI	PTGPGYQVPP	-----	AARVXAGST	LPVAAV
NP_295913	GMDLEQTVQ	LSKLLK-Y	BGVVDLDS	GGLTAAQ	-----	QI	EVGPGYQVPP	-----	AAVVSRAETE	ISVMAV
AF320254	GNTADDAVA	IARLFK-B	AGADIDCSS	GQVWKGQ	-----	OP	VYGRMYQTFF	-----	ADRIRNEVG	PTLAVG
OYE family	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
AF4875	EL-VQPEY	LIA	OM	RRLDVAYLHL	-----	DE	EKPHDPNHE	-----	VFVRVWG-Q	SS-PILLA
AF4961	EQR-VETWTF	LCESSLKKAHF	NLSYVSF	IEPRYE	-----	QIFSYEBKO	NFLRSWG	-----	LSVDVLSFR	KIPGTTPPFS
Ca2460	BE	ILHSY	ILQQLQORAD	NGQOLAYVSL	-----	IFDASL	EDQGRSNEF	-----	AKYKWK	-----
Ne4452	DLIP-OFED	VIRKIN	GFCLAYLHL	TQSRVAGN	-----	KDVOP	EEDSE-NLAP	-----	AAKLMDG	-----
ScOYE1	ETGIVAQVAY	VAGELEKRAK	AGKRLAPVHL	VEPRVTNP	-----	PLTEGE	GEYEGGSNDP	-----	VYSIWK	-----
ScOYE2	ETGIVAQVAY	VAGELEKRAK	AGKRLAPVHL	VEPRVTNP	-----	PLTEGE	GEYEGGSNDP	-----	AYSIMWK	-----
ScOYE3	EPGLIAQVSY	VAGELEKRAK	AGKRLAPVHL	VEPRVTNP	-----	SLVEGE	GEYSEGTNDP	-----	AYSIMWK	-----
A36990	BE	ILHSY	ILQQLQORAD	NGQOLAYVSL	-----	IYDVSL	KDQQGRSNEF	-----	AYKIMWK	-----
	501	511	521	531	541	551	561	571	581	591
	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 3	NGKO--AN	QILEEQD	-----	-----	-----	IDVALVG	RGFOKDPGLA	-----	WTFPAQLGV	-----
SEQ 6	SAHLANS	LLEKGG	-----	-----	-----	LDLVLVG	RGFOKDPGLV	-----	WMADELNV	-----
SEQ 8	ADEATAAAM	LSGPEPK	-----	-----	-----	ADAILIA	RQFLREPEWV	-----	FSTARKLGV	-----
SEQ 10	NGKO--AN	KLLEEDG	-----	-----	-----	LDLVALVG	RGFOKDPGLA	-----	WTFPAQLGV	-----
SEQ 12	RDPELLN	KYLEEGT	-----	-----	-----	FDLALIG	RGFLRNPGLV	-----	WEPADKLG	-----
SEQ 14	RDIFKLD	EPIANGD	-----	-----	-----	FDIALIG	KGFLKNTGLI	-----	SRIADQLQA	-----
SEQ 16	E-DGRVTIQR	ENGAKTR	-----	-----	-----	ADMVLVA	ROFLKEPEFV	-----	LTVADELGV	-----
SEQ 19	TRQOME	AALSDDD	-----	-----	-----	CDMIGIG	RPALINPSLP	-----	ANLILNPEV	-----
SEQ 22	TADI--AR	DVDEQGAEEK	VAEAKQTHDT	IEVVSESHGG	-----	RTKADILVIA	RQFLREPEFV	-----	LRTABNLGV	-----
SEQ 24	DGLFTTAN	EILESCK	-----	-----	-----	ADVTFA	REFLRNPSLV	-----	LDANQLGEB	-----
SEQ 27	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 30	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 33	TGHLV--AE	EVLQSG	-----	-----	-----	IDIVRAG	RMFOONPGLV	-----	RAFANELGV	-----
SEQ 35	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 38	ETGIVAQVAY	VAGELEKRAK	AGKRLAPVHL	VEPRVTNP	-----	-----	-----	-----	-----	-----
SEQ 40	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 42	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 44	VGAM-VDA	LQGVQD	-----	-----	-----	IGIG	RAAGSEPDIA	-----	KDILACRVSS	-----
SEQ 83	IGTLV--AE	EILAGG	-----	-----	-----	ED	DTPLDLVAGS	-----	RLFOKNTGLV	-----
SEQ 85	VGMM--EG	SYDSFNG	-----	-----	-----	QDRSOG	KLAESQISQSG	-----	ECDVALLAR	-----
Bacteria	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
T44612	POLAE	QALQANO	-----	-----	-----	LDLVSVG	RAHLADPHNA	-----	YFAAKELGV	-----
NP_625402	EPQ	KILANGE	-----	-----	-----	ADAVLLG	RELLRNPSMA	-----	QHAARELGV	-----
NP_295913	TGA--QAE	AIIQAGD	-----	-----	-----	ADLIALG	RPFLRDPHNA	-----	QRAARELGL	-----
AF320254	AD--HAN	SITAAGR	-----	-----	-----	ADLCAIA	RPHLADPAWT	-----	LHEAAKIGF	-----
OYE family	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
AF4875	AASAEVTEQ	MAATYTF	-----	-----	-----	NVAIATP	RYFISTPDLP	-----	FRVMAGTQL	-----
AF4961	AGQDQSNW	GVLSECB	-----	-----	-----	YDALLYG	RYFTSNPDIV	-----	ERLKKGIFF	-----
Ca2460	YDAPEFTKL	LHDLND	-----	-----	-----	RTIVGFA	RPFTSNPDIV	-----	EKLKLGKPL	-----
Ne4452	PETAK-HLV	DREFFEK	-----	-----	-----	DVATFC	RHFISTPDLP	-----	FRIKBGIEL	-----
ScOYE1	LHP--EVV	REEVKDK	-----	-----	-----	RTLIGYG	RPFTSNPDIV	-----	DRLEKGLPL	-----
ScOYE2	LHP--EVV	REEVKDK	-----	-----	-----	RTLIGYG	RPFTSNPDIV	-----	DRLEKGLPL	-----
ScOYE3	LHP--EVV	REEVKDK	-----	-----	-----	RTLIGYG	RPFTSNPDIV	-----	YRLEEGPL	-----
A36990	YDAPEFTKL	INDLND	-----	-----	-----	RSIIGFS	RPFTSNPDIV	-----	EKLKLGKPL	-----

	601	611	621	631
SEQ 3	-----	-----	-----	-----
SEQ 6	-----	-----	-----	-----
SEQ 8	-----	-----	-----	-----
SEQ 10	-----	-----	-----	-----
SEQ 12	-----	-----	-----	-----
SEQ 14	-----	-----	-----	-----
SEQ 16	-----	-----	-----	-----
SEQ 19	-----	-----	-----	-----
SEQ 22	-----	-----	-----	-----
SEQ 24	-----	-----	-----	-----
SEQ 27	-----	-----	-----	-----
SEQ 30	-----	-----	-----	-----
SEQ 33	-----	-----	-----	-----
SEQ 35	-----	-----	-----	-----
SEQ 38	-----	-----	-----	-----
SEQ 40	-----	-----	-----	-----
SEQ 42	-----	-----	-----	-----
SEQ 44	-----	-----	-----	-----
SEQ 83	-----	-----	-----	-----
SEQ 85	-----	-----	-----	-----
Bacteria	-----	-----	-----	-----
T44612	-----	-----	-----	-----
NP_625402	-----	-----	-----	-----
NP_295913	-----	-----	-----	-----
AF320254	-----	-----	-----	-----
OYE family	-----	-----	-----	-----
AF4875	-----	-----	-----	-----
AF4961	-----	-----	-----	-----
Ca2460	-----	-----	-----	-----
Nc4452	-----	-----	-----	-----
ScOYE1	-----	-----	-----	-----
ScOYE2	-----	-----	-----	-----
ScOYE3	-----	-----	-----	-----
A36990	-----	-----	-----	-----

Figure 1. A multiple alignment of the 2031 OR amino acid sequence from *A. fumigatus* (SEQ ID No3) along with related 2031 ORs from other fungi and bacteria (see Example 4) and OYEs. Regions 1-11, marked with * or #, refer to amino acids conserved between ORs but not OYEs.

Fungal 2031 ORs are given by the following SEQ ID No.: *A. fumigatus*, SEQ ID Nos. 3, 6 and 8; *A. nidulans*, SEQ ID No. 10; *C. albicans* SEQ ID Nos. 12 and 14; *N. crassa*, SEQ ID Nos. 16 and 19; *M. grisea* SEQ ID Nos. 22 and 44; *S. pombe* SEQ ID No. 24 (NP_595868); *C. trifolii* SEQ ID No. 27; *F. sporotrichioides* SEQ ID Nos. 30, 33 and 35; *F. graminearum* SEQ ID Nos. 38 and 83; *M. graminicola* SEQ ID Nos. 40 and 42; *U. maydis* SEQ ID No 85.

Bacterial ORs resembling 2031 are:

T44612 (*Pseudomonas putida*), SEQ ID No. 86; NP_625402 (*Streptomyces coelicolor*), SEQ ID No. 87; NP_295913 (*Deinococcus radiodurans*), SEQ ID No. 88; AF320254 (*Azoarcus evansii*, SEQ ID No. 89).

Fungal ORs similar to the Old Yellow Enzyme family (originally identified in *S. cerevisiae*):

A. fumigatus, Af4875 and Af4961, SEQ ID Nos. 90 and 91 respectively; *C. albicans*, Ca2460 and A36990, SEQ ID Nos. 92 and 93 respectively; *N. crassa*, Nc4452, SEQ ID No. 94; *S. cerevisiae*, OYE1, OYE2 and OYE3, SEQ ID Nos. 95-97 respectively.

Details of the sequence searches that identified the ORs other than SEQ ID No. 3, and methods for the construction of multiple alignments are given in Example 4 hereinafter.

	1	11	21	31	41	51	61	71	81	91
SEQ 1	GTTCGACGTC	ATTGCCACGT	TTGACCCAA	GGGCAGACGC	CATGTCGCGG	AGCGATCGCC	GCGATATGCC	TCGAATTGTC	GCCATTCGGC	ATCCAGTTTC
SEQ 2	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 4	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 5	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 7	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 9	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 11	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 13	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 15	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 17	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 18	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 20	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 21	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 23	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 25	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 26	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 28	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 29	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 32	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 34	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 36	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 37	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 39	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 41	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 43	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 82	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 84	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
	101	111	121	131	141	151	161	171	181	191
SEQ 1	CAGTGCCCTT	CCCCGAATGA	CTGTCTCCAC	TAITCGGCAA	GATTGTAAAT	CAAGCCTGAA	GAAGCGGAGC	AATTCTTGGA	AGTCGTATGT	TCTACTGATT
SEQ 2	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 4	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 5	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 7	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 9	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 11	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 13	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 15	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 17	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 18	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 20	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 21	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 23	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 25	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 26	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 28	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 29	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 32	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 34	---AGGAAG	TTGCAATGCA	CTTGTAAGTGA	CAGGGCGTCG	TGTAAATTTT	ATAAATACCT	ATACTTGTIT	GTTCACTTCT	ATGCTACTCA	TATCAATCCG
SEQ 36	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 37	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 39	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 41	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 43	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 82	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 84	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
	201	211	221	231	241	251	261	271	281	291
SEQ 1	TCTGTGCCTG	GCGCAGACGG	GTATATAAAT	AAAGATCACC	GCACCGAGGA	GTTTCTTACC	AACCCATCAA	TAACCATCCA	CAATCTCCTA	CAACAAAAAT
SEQ 2	TCTGTGCCTG	GCGCAGACGG	GTATATAAAT	AAAGATCACC	GCACCGAGGA	GTTTCTTACC	AACCCATCAA	TAACCATCCA	CAATCTCCTA	CAACAAAAAT
SEQ 4	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 5	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 7	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 9	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 11	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 13	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 15	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 17	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 18	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 20	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 21	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 23	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 25	TGACCTCTCT	CTTGACAACA	AAGCCGGCCA	TCCTCGCCGA	CGATTGCTCT	TACCCCGCA	TAGTCACACT	CGCACGTCCG	TTCTCCCACT	GTCAAAACAGA
SEQ 26	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 28	TGCTGTAGAT	GTGGTTGAAT	TGGTATATTA	GACCGGAGTA	CTCTATATGC	GAGAGACTAT	ACATTGAAGT	TGCCAACGTT	CTTCCAGATT	GATTAAATCAT
SEQ 29	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 32	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 34	AGAAGATCAA	ACAGTCCCTT	ATACACACTT	GTCAAGACCT	ATCTATTATT	TCAAAAATCA	GCAATATGGC	TGAGACAATG	CCTAAGTGTG	AGGCAAAATGG
SEQ 36	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 37	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 39	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 41	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 43	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 82	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 84	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

	301	311	321	331	341	351	361	371	381	391
SEQ 1	GA	GA	TT	CG	CC	CT	GA	GC	CA	CC
SEQ 2	AG	AT	TC	GG	CT	CG	AC	CG	AG	CA
SEQ 3	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 4	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 5	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 6	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 7	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 8	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 9	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 10	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 11	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 12	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 13	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 14	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 15	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 16	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 17	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 18	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 19	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 20	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 21	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 22	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 23	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 24	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 25	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 26	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 27	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 28	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 29	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 30	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 31	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 32	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 33	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 34	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 35	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 36	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 37	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 38	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 39	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 40	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 41	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 42	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 43	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 44	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 45	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 46	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 47	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 48	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 49	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 50	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 51	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 52	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 53	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 54	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 55	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 56	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 57	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 58	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 59	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 60	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 61	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 62	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 63	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 64	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 65	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 66	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 67	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 68	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 69	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 70	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 71	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 72	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 73	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 74	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 75	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 76	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 77	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 78	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 79	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 80	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 81	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 82	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 83	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA
SEQ 84	TT	TC	CT	GG	CT	CG	AC	CG	AG	CA

```

601      611      621      631      641      651      661      671      681      691
--2-----3-----
*****
SEQ 1  CAATACTCCG CC----- --CAGGACG GCCACATGAC CGAC----- TACCACATCG CCCATCTGGG TGGGATCGCC CAACCGGGAC
SEQ 2  CAATACTCCG CC----- --CAGGACG GCCACATGAC CGAC----- TACCACATCG CCCATCTGGG TGGGATCGCC CAACCGGGAC
SEQ 4  CAATACTCCG CC----- --GACGATG GACACATGAC TCCC----- TGGCATATGG CACATCTTGG AGGGATTGCC CAGCGAGGGC
SEQ 5  CAATACTCCG CC----- --GACGATG GACACATGAC TCCC----- TGGCATATGG CACATCTTGG AGGGATTGCC CAGCGAGGGC
SEQ 7  ATGTACTCCT GCGAGTCGGA CCCGTCGTCT CCCACGCTCG CGCGCCTAAC AAAC----- TACCACCTGG CGCATCTGGG CCACCTCGCC CTCAAAGGCG
SEQ 9  CAGTACTCCG CA----- --GAAGACG GCCACATGAC AGAC----- TACCACATCG CGCACTTGGG AGGTATTGCC CAGCGCGGCC
SEQ 11 CAATATTCTG CT----- GATTATAATT TTGAAGCAAC TCCA----- TACCATTAA TCCATTATGG TTCAITAGTG AATCTGGGG
SEQ 13 ATGTATTCTG CG----- --TCA CCAACTGACA ATCAAGCCAC TCTG----- TTTCATTGTG TCCATTATGG ATCAITGTCT GTACGTGGAC
SEQ 15 ACCTACTCAT CC----- --GACGATG GACACATGAC CGAC----- TGGCACTTGG TCCACCTGGG CTCTTTCGCC CTCGCCGGTG
SEQ 17 GAACAAATGG GC----- --TTCGCA ACCACTTGGC CAAC----- CCCGAACCTG CCGCGCTCTA CGCCACCTGG GCCCGGGCG
SEQ 18 GAACAAATGG GC----- --TTCGCA ACCACTTGGC CAAC----- CCCGAACCTG CCGCGCTCTA CGCCACCTGG GCCCGGGCG
SEQ 20 ACCTACTCAG CC----- --GACGATG GCCACCTGAC CGAC----- TTCCACTTGG TGCACCTGGG CCAGTTTCGCC CTGCAACGGCA
SEQ 21 ACCTACTCAG CC----- --GACGATG GCCACCTGAC CGAC----- TTCCACTTGG TGCACCTGGG CCAGTTTCGCC CTGCAACGGCA
SEQ 23 ACTTATTCCG CT----- GACCAAGAGG GGCATTGAC AGAT----- TTTCACCTAG TACATCTTGG AGCGATTGGGA ATGCGTGGGG
SEQ 25 CAGTACTCCG CC----- --GACAAATG GCCACGGAC CGAC----- TACCACCTCG TCCACCTGGG CCAAGTTCGCC CTGCAACGGC
SEQ 26 CAGTACTCCG CC----- --GACAAATG GCCACGGAC CGAC----- TACCACCTCG TCCACCTGGG CCAAGTTCGCC CTGCAACGGC
SEQ 28 CAATACAGTG CC----- --AAAGATG GTTATGCCAC TGAT----- TGGCACTTGA CTCACCTCGG GGGAAATATC CAAAGAGGCC
SEQ 29 CAATACTCCG CC----- --AAAGATG GTTATGCCAC TGAT----- TGGCACTTGA CTCACCTCGG GGGAAATATC CAAAGAGGCC
SEQ 32 -----
SEQ 34 CAATATTCTG CA----- --AAGGATG GTGTCTGAC CCCC----- TGGCACAAC AACACCTGGG CAGCTTCGCA GCACGAGGTC
SEQ 36 CAATACTCCG CC----- --AAAGATG GATATGCTAC TGAT----- TGGCACTTGA CTCATCTCGG AGGCATTATC CAACGAGGCC
SEQ 37 CAATACTCCG CC----- --AAAGATG GATATGCTAC TGAT----- TGGCACTTGA CTCATCTCGG AGGCATTATC CAACGAGGCC
SEQ 39 CAGTACTCTG CT----- --CCGACG GACACTACAC AATG----- TGGCATCACA CCCACATGGG CGGCATATC CAACCGGGTC
SEQ 41 -----
SEQ 43 GAGGGCCCTG CG-- --ACGTT TGACGAGGCG GACCCGTCGA AGCGCGGACG CCCGACGGAG CAGCTGGTGG AGCTGTACCG CGCCTGGGGC CAGGGCGAGT
SEQ 82 CAATACAGTG CC----- --CGTAGC GCTTTCACGA GCCT----- TGGCACTTGG CCCACTCGG CGCACTGGCC CAGCTGGGCC
SEQ 84 CAGTACTCTG CG----- --AACAAATG GTCTTCTCTAC TCCG----- TACCACATTG CGCATTTGGG ATCGTTTGGC CTGCAACGGTG

701      711      721      731      741      751      761      771      781      791
-----4-----
*****
SEQ 1  CCGGCCCTGAT GCTGATTGAG GCGACCGCCG TCCAGCCCGA A-- --GGCCGC ATCACCCCTC AGGATGTGCG TCTGTGGAAG GACTCC----- CA
SEQ 2  CCGGCCCTGAT GCTGATTGAG GCGACCGCCG TCCAGCCCGA A-- --GGCCGC ATCACCCCTC AGGATGTGCG TCTGTGGAAG GACTCC----- CA
SEQ 4  CAGGATTTCTT GATGGTCGAG GCAACAGCAG TCGAACCCGA A-- --GGCAGG ATCACCCCGC AGGACCTGGG ACTATGGAAG GACTCG----- CA
SEQ 5  CAGGATTTCTT GATGGTCGAG GCAACAGCAG TCGAACCCGA A-- --GGCAGG ATCACCCCGC AGGACCTGGG ACTATGGAAG GACTCG----- CA
SEQ 7  CAGGCCCTCTG CTTTCATCGAA GCCACCGCCG TCGACCCCAA A-- --GGCGCG ATCTCCCCCA ACGACTCGGG CCTCTGGCAG GACGCGACCA CCTCGGAACA
SEQ 9  CCGGCTCTCAT GATGATCGAG GCAACCTCCG TCTCATCGAA A-- --GGCAGA ATCACCGCCG AGGACGTGCG TTTATGGAAG GACTCG----- CA
SEQ 11 CAGGATTTCTT CATTGTTAGAG AGCAACGGCTG TTTCTCCTGA A-- --GGTGGG TTATCACCTC ATGATTATAG AATCTGGAAG GATGAA----- CA
SEQ 13 CAGCATTAAT CATTGTTAGAG AGTATCTTTG TGTCGCAAG T-- --TCCGGA TTATCCATTG ATGATTATAG TCTTTGGAAT GATGAT----- CA
SEQ 15 TCCCCCTCAC CATCTTCGAG GCCACCGCCG TCCCTCCCAA C-- --GGCCGC ATCACCCCGC AGTGTCTGCG TCTCTGGCAG GACTCC----- CA
SEQ 17 ACTGGGGCCT GATTCTCACC GGCAACGTCC AAGTCGACCA CGCGCACAA GCGGACGCCC ACGACATCAG CCCCACCAAC CCGCGGACCA CCGCCGAGCA
SEQ 18 ACTGGGGCCT GATTCTCACC GGCAACGTCC AAGTCGACCA CGCGCACAA GCGGACGCCC ACGACATCAG CCCCACCAAC CCGCGGACCA CCGCCGAGCA
SEQ 20 CGGCCCTGAC CATTTGTCGAG GCCACATCCG TCACGCCCAA C-- --GGACGC ATCTCGCCCG AGGACAGCGG CCTGTGGCAA GACAGC----- CA
SEQ 21 CGGCCCTGAC CATTTGTCGAG GCCACATCCG TCACGCCCAA C-- --GGACGC ATCTCGCCCG AGGACAGCGG CCTGTGGCAA GACAGC----- CA
SEQ 23 CTGGCCTTGT AATGGTAGAA GCGACAGCGG TTTCGCCAGA G-- --GGACGA ATTTACCTTA ATGATTGATG ATTAATGATG GAGTCG----- CA
SEQ 25 CGGCCCTGTC CATGGTCGAG GCCACCGCCG TCGAGGCTCG T-- --GGCCGC ATCTCCCCCG AGGATGTGCG TTTGTGGCAG GACTCG----- CA
SEQ 26 CGGCCCTGTC CATGGTCGAG GCCACCGCCG TCGAGGCTCG T-- --GGCCGC ATCTCCCCCG AGGATGTGCG TTTGTGGCAG GACTCG----- CA
SEQ 28 CCGGATTGTC CATGGTCGAG GCTACCGCTG TACAAAACCA C-- --GGTCGC ATCACACCTC AGGATGTTGG TCTGTGGGAA GACGGC----- CA
SEQ 29 CCGGATTGTC CATGGTCGAG GCTACCGCTG TACAAAACCA C-- --GGTCGC ATCACACCTC AGGATGTTGG TCTGTGGGAA GACGGC----- CA
SEQ 32 -----
SEQ 34 CGGGTCTCAT TGTTCACAGAA GTCAACGCG TTTCCACCAGA G-- --GGACGA ATCAGTCTCG AGGATGCGAG CATCTACGAT GATGGG----- CA
SEQ 36 CGGGTCTCAT TGTTCACAGAA GTCAACGCG TTTCCACCAGA C-- --GGTCGC ATCAGGCTC AGGACGTTGG TCTCTGGGAA GATGGA----- CA
SEQ 37 CGGGTCTCAT TGTTCACAGAA GTCAACGCG TTTCCACCAGA C-- --GGTCGC ATCAGGCTC AGGACGTTGG TCTCTGGGAA GATGGA----- CA
SEQ 39 CCGGACTGTC CATGGTCGAG GCGACAGCGG TGACTCTCTA A-- --GGTCGC ATCACGCTCG AAGACGTCGG TATCTGGCAA GATTCT----- CA
SEQ 41 -----
SEQ 43 GGGGCCAGAT CCAGAGCGGG AACGTCTAGA TCGACCCGGA GCACCTCGAG GCCCGGGGCA ACATGTTGGT GCGCGCGGAC GCGGAGCCCT CCGGGGAGCG
SEQ 82 CTGCGCTCAT CATGCTAGAA GCTACCGCAG TTCAAGCAGG T-- --GGCCGT ATCACACCTG AAGATTCTGG CATCTGGCTA GACTCT----- CA
SEQ 84 TGGGAAACGT CATGGTCGAA GCATCTGGTG TTGAGCCAGA G-- --GGGAGG ATCACGCTC AGGACCTGGG TATTTGGTGC GATACG----- CA

801      811      821      831      841      851      861      871      881      891
-----5-----
*****
SEQ 1  GATCGCCCGG --- --ATGCGCC GGGTCATCGA CTTCTGTGAC AGCCAGGGG- CAGAAGATCG GCGTG----- --CAGCTT GCCCATGCGG GCCGGAAGGC
SEQ 2  GATCGCCCGG --- --ATGCGCC GGGTCATCGA CTTCTGTGAC AGCCAGGGG- CAGAAGATCG GCGTG----- --CAGCTT GCCCATGCGG GCCGGAAGGC
SEQ 4  GATTGAGCCA --- --TTGAGCC GCGTGATCGA GTTTGTCCAC AGTCAGAAC- CAGCTTATCG GCGTG----- --CAGATC GCACACGCGG GTCCGAAGGC
SEQ 5  GATTGAGCCA --- --TTGAGCC GCGTGATCGA GTTTGTCCAC AGTCAGAAC- CAGCTTATCG GCGTG----- --CAGATC GCACACGCGG GTCCGAAGGC
SEQ 7  ATTCTTGGGG --- --CTGAAGC GGGTCGTCGA GTTCTGTGAC GCACAGGGG- CAGCAAGGTC GCGTG----- --CAGCTT GCGCATGCGG GCCGGAAGGC
SEQ 9  GATTGCGCCC --- --ATGAAGC GCGTCATCGA CTTCTGTGAC TCGCAGTCC- CAGAAGATTG GCGTG----- --CAGATT GCCCAGCGCG GCCGGAAGGC
SEQ 11 AGCAGAGAAA --- --TTGAAGC CAATTGTGCGA TTACGCTCAT TCTCAAAG- CAATTAAATTG CCATC----- --CAATTG GCGCATGGTG GTAGAAAAGC
SEQ 13 AGCTCACAGT --- --TTACGGA AAATGTGTGA TTTTATTAT GATCAAAG- GGAATTTGTCT GTATA----- --CAATTG AATCACGCTG GCGGAAAGAT
SEQ 15 GATTGCGCCC --- --CTCAAGC GCATCGTCGA CTACATCCAC TCCCAGGGG- CAGAAAGGCGG GTATC----- --CAGCTT GCCCAGCGCG GCCGGAAGGC
SEQ 17 GACCGTCACG GCCTTCAAGG CCTGGGCGGA CGCCGCGCGC CTGAATGGG- CAGTCCAAAA CGCTGTGGT CTGTGAGATC AACCAACCTG GTCCGCAAGG
SEQ 18 GACCGTCACG GCCTTCAAGG CCTGGGCGGA CGCCGCGCGC CTGAATGGG- CAGTCCAAAA CGCTGTGGT CTGTGAGATC AACCAACCTG GTCCGCAAGG
SEQ 20 GATCGCTCTC --- --CTGCGCC GCATCGTCGA CTACGTGAC AGCCAGGGG- CAAAAGATCG CCATC----- --CAACTG GCTCATGCGG GCCGCAAGGC
SEQ 21 GATCGCTCTC --- --CTGCGCC GCATCGTCGA CTACGTGAC AGCCAGGGG- CAAAAGATCG CCATC----- --CAACTG GCTCATGCGG GCCGCAAGGC
SEQ 23 AATGAAGCGG --- --TTACGGA GAATGTGTGA ATTTGTCTAT TCGCAAAAT- CAAAAAATTG GGATT----- --CAATTG GCGCATGCTG GTAGAAAAGC
SEQ 25 GATTGCGCGG --- --CTGAAGC GCATCGTCGA CTTTATCCAC TCGCAGAAC- CAGGTTCGCGG CCATC----- --CAGCTC GCCCAGCGCG GTCCGAAGGC
SEQ 26 GATTGCGCGG --- --CTGAAGC GCATCGTCGA CTTTATCCAC TCGCAGAAC- CAGGTTCGCGG CCATC----- --CAGCTC GCCCAGCGCG GTCCGAAGGC
SEQ 28 GATGAGCGCT --- --CTGAAGC GCATCAACAC TTTCGCGCAC AGTCAGAGG- CAGAAAATTG GTATC----- --CAGCTG TCGCATGCGG GTCCGAAGGC
SEQ 29 GATGAGCGCT --- --CTGAAGC GCATCAACAC TTTCGCGCAC AGTCAGAGG- CAGAAAATTG GTATC----- --CAGCTG TCGCATGCGG GTCCGAAGGC
SEQ 32 -----
SEQ 34 GCTTGGACCT --- --CTCGGGG ATATTGTGGA CTTTGTACAC AGCCAGGGG- GCCAAGATTG CTATT----- --CAGATA GGTTCATGCTG GGAGAAAAGC
SEQ 36 AATCGAGCCC T-- --TTGAAGC GCATCACTAC TTTTGGCCAC AGCCAAAGG- CAGAAGATTG GTAT----- --TCAGCTC TCGCAGCTG GTCTGTAAGGC
SEQ 37 AATCGAGCCC --- --TTGAAGC GCATCACTAC TTTTGGCCAC AGCCAAAGG- CAGAAGATTG GTAT----- --TCAGCTC TCGCAGCTG GTCTGTAAGGC
SEQ 39 GATCGAGCCT C-- --TTGCCAA GGTCTGTC- GA GTTTGCCAC TCCCAGAAC- CAGAAGATCA TGATT----- --CAGTTG GCGCATGCGG GCCGCAAGGC
SEQ 41 -----
SEQ 43 CTTGCAATG TTTTCAAGC TCGCGCGCG CGCCAAAGG- CACGGCAGG- CTC-ATCGTC GCG----- --CAGGTC GGACACCCCG GTCCGCAAGC
SEQ 82 TGTGAGGGA --- --CTGCGAA AGCAGCTGGA TTTTGGCCAC GCCTTATCG GTATC----- --CAGATT GGCCATGCTG GTCCGCAAGC
SEQ 84 TCGGGATGCA --- --CACAAGG CGCTGGTGTG GGTGCTCAAG TCCTTCAAG- GATGGTCTG GTGTA----- --GGGCTG CAACTGGCGC ATGCGGGAAG

```

	901	911	921	931	941	951	961	971	981	991
SEQ 1	*****	*****	*****	*****	*****	*****	*****	*****	*****	*****
SEQ 2	CACCAACGTT	GGCCCTTGA	TCTCA	-----	-----	-----	-TTCTCGGCC	ATCGCGACGG	AGAAGTGGG	CGGATGGCCG
SEQ 3	CACCAACGTT	GGCCCTTGA	TCTCA	-----	-----	-----	-TTCTCGGCC	ATCGCGACGG	AGAAGTGGG	CGGATGGCCG
SEQ 4	CAGCAACGTC	GGCCCATGGC	TCTCG	-----	-----	-----	-GCCAACGAT	ACCGCCTCCG	AGAAGTGGG	CGGCTGGCCA
SEQ 5	GAGCAACGTC	GGCCCATGGC	TCTCG	-----	-----	-----	-GCCAACGAT	ACCGCCTCCG	AGAAGTGGG	CGGCTGGCCA
SEQ 7	GAGTGGCGTG	GGCCCGTGGC	TGCGG	-----	-----	-----	-----	-----	-----	-----
SEQ 9	TTCTGACATC	GGCCCTTGGC	TCATG	-----	-----	-----	-----	-----	-----	-----
SEQ 11	TTCTGATCAG	CCCTTATTTT	TGCAC	-----	-----	-----	-----	-----	-----	-----
SEQ 13	TGTTGAAGGG	GTACCAATTC	AACAA	-----	-----	-----	-----	-----	-----	-----
SEQ 15	CTCCACCAAG	GGCCCTTGGC	ACTAC	-----	-----	-----	-----	-----	-----	-----
SEQ 17	TCCGATGGGC	GGGGGACGC	GGGGA	-----	-----	-----	-----	-----	-----	-----
SEQ 18	TCCGATGGGC	GGGGGACGC	GGGGA	-----	-----	-----	-----	-----	-----	-----
SEQ 20	CAGCACAAAG	GGCCCTTGGC	ACGACTCCTT	CACCCCAAGC	GGCGAGTATA	AGCCGAGAGA	GGGCTTACAG	GTGCTCGGAC	CCGAGTATGG	CGGCTGGCCT
SEQ 21	CAGCACAAAG	GGCCCTTGGC	ACGACTCCTT	CACCCCAAGC	GGCGAGTATA	AGCCGAGAGA	GGGCTTACAG	GTGCTCGGAC	CCGAGTATGG	CGGCTGGCCT
SEQ 23	TAGCAACACT	GCTCTTATC	GAGGA	-----	-----	-----	-----	-----	-----	-----
SEQ 25	TAGCAACCTG	GCACCGTGA	TCACC	-----	-----	-----	-----	-----	-----	-----
SEQ 26	TAGCAACCTG	GCACCGTGA	TCACC	-----	-----	-----	-----	-----	-----	-----
SEQ 28	CAGTTGCGTA	TCTCCCTGGC	TAAGC	-----	-----	-----	-----	-----	-----	-----
SEQ 29	CAGTTGCGTA	TCTCCCTGGC	TAAGC	-----	-----	-----	-----	-----	-----	-----
SEQ 32	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 34	GAGCACAGTC	GTACCGTGGC	TGGAC	-----	-----	-----	-----	-----	-----	-----
SEQ 36	TAGTTGTGTA	TCTCCGTGGT	TGAGC	-----	-----	-----	-----	-----	-----	-----
SEQ 37	TAGTTGTGTA	TCTCCGTGGT	TGAGC	-----	-----	-----	-----	-----	-----	-----
SEQ 39	GAGCACTGTG	GCACCATGGT	TAAGC	-----	-----	-----	-----	-----	-----	-----
SEQ 41	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 43	CGCGCGCAGC	GTCCAGCAGC	ACCCC	-----	-----	-----	-----	-----	-----	-----
SEQ 82	CTCCTGCGTT	GCTCCTGGT	TAGAC	-----	-----	-----	-----	-----	-----	-----
SEQ 84	GAAGGCCCTG	GACTGGTCAC	CTTTC	-----	-----	-----	-----	-----	-----	-----
	1001	1011	1021	1031	1041	1051	1061	1071	1081	1091
SEQ 1	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 2	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 4	GAC - CGCGTC	AAAGGGCCCG	GGCATATC	-----	-----	-----	-----	-----	-----	-----
SEQ 5	GAC - CGCGTC	AAAGGGCCCG	GGCATATC	-----	-----	-----	-----	-----	-----	-----
SEQ 7	GAC - CGCGTC	AAAGGGCCCG	GGCATATC	-----	-----	-----	-----	-----	-----	-----
SEQ 9	GAT - CGTGTG	ATCGGCCCTG	CCACCGTG	-----	-----	-----	-----	-----	-----	-----
SEQ 11	GAC - AAGGCA	GTGCTCCTT	CTGCTATG	-----	-----	-----	-----	-----	-----	-----
SEQ 13	GAA - CATGTG	GTGGGGCCAT	CTACTGAG	-----	-----	-----	-----	-----	-----	-----
SEQ 15	GAG - AACGCT	TGGGCCCCCA	GCGCCATC	-----	-----	-----	-----	-----	-----	-----
SEQ 17	TTT - GTGCCT	CGCTTGTGTG	CGAAAGTG	-----	-----	-----	-----	-----	-----	-----
SEQ 18	TTT - GTGCCT	CGCTTGTGTG	CGAAAGTG	-----	-----	-----	-----	-----	-----	-----
SEQ 20	GAT - GACGTC	TGGGCCCCCA	GCGCCATC	-----	-----	-----	-----	-----	-----	-----
SEQ 21	GAT - GACGTC	TGGGCCCCCA	GCGCCATC	-----	-----	-----	-----	-----	-----	-----
SEQ 23	AAT - GATGTT	TATGGACCAA	ATGAAGAC	-----	-----	-----	-----	-----	-----	-----
SEQ 25	GAC - GACGTT	GTGGCTCCCA	GCGCGATT	-----	-----	-----	-----	-----	-----	-----
SEQ 26	GAC - GACGTT	GTGGCTCCCA	GCGCGATT	-----	-----	-----	-----	-----	-----	-----
SEQ 28	GAC - AATATC	GTGCTCCCT	CGGCCATC	-----	-----	-----	-----	-----	-----	-----
SEQ 29	GAC - AATATC	GTGCTCCCT	CGGCCATC	-----	-----	-----	-----	-----	-----	-----
SEQ 32	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 34	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 36	GAC - AACATT	GTGCTCCTT	CTGCCATC	-----	-----	-----	-----	-----	-----	-----
SEQ 37	GAC - AACATT	GTGCTCCTT	CTGCCATC	-----	-----	-----	-----	-----	-----	-----
SEQ 39	GAG - GATGTC	TGGGGCCCA	GTGCGATT	-----	-----	-----	-----	-----	-----	-----
SEQ 41	GAG - GATGTT	GTGGGTCCCT	CGGGTGGGGA	GGACTTTACG	TGGGATGAGA	GGTCTCGAG	CGACCTTAGT	GGAGGCTACT	ATGCGCCGAG	AGAGTTGTGC
SEQ 43	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 82	GAT - GCGTGT	GTGCGACCTA	GCAACGAG	-----	-----	-----	-----	-----	-----	-----
SEQ 84	GAT - GCGTGT	GTGCGTCTT	CGGCCATC	-----	-----	-----	-----	-----	-----	-----
	1101	1111	1121	1131	1141	1151	1161	1171	1181	1191
SEQ 1	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 2	CTGGATGA - G	ATCGAGCAGT	TCAAGAAGGA	CTGGGTGGCG	GCCACGAAGC	GGGCCATCGC	CG - - - CCGGT	GGGACTTTTG	TCGAGATTCA	CAATGCGCAT
SEQ 3	CTGGATGA - G	ATCGAGCAGT	TCAAGAAGGA	CTGGGTGGCG	GCCACGAAGC	GGGCCATCGC	CG - - - CCGGT	GGGACTTTTG	TCGAGATTCA	CAATGCGCAT
SEQ 4	AAGCAGGA - T	ATCGAGGATC	TGAAGACCGC	CTGGGTGGCC	GCTGTCAAAC	GGGCTGTATA	GG - - - CCGGA	GGGACTTTTA	TCGAGATCCA	CAATGCGCAT
SEQ 5	AAGCAGGA - T	ATCGAGGATC	TGAAGACCGC	CTGGGTGGCC	GCTGTCAAAC	GGGCTGTATA	GG - - - CCGGA	GGGACTTTTA	TCGAGATCCA	CAATGCGCAT
SEQ 7	ACGGCCGA - G	GTCCGTCAGG	TGGTGGCGGC	GTGTCGAAG	AGCGCGCGGC	TAGCGGTGCA	GG - - - CTGGG	GTGGATGTTA	TCGAGATCCA	TGGGCGGCAT
SEQ 9	AAGCAGGA - C	ATCGAGCAGT	TCAAGACCGA	CTGGTGTGAT	GCGTGCAGAG	GGGCCATTGC	CG - - - CTGGC	GGGACTTTCA	TCGAGATCCA	CAATGCCCAT
SEQ 11	AAAGATGA - A	ATCAAACTGT	TGTTAAAGGA	TTTGTGCTG	GAGCTGTGTA	AATCAGTGGC	TTTGTGTCAG	TTTGAATGCA	TTTGAATGCA	TTTGTGCTAT
SEQ 13	GTAAATGA - A	ATAAATTCAT	TGTTGAAGGA	CTTTGCCAAT	GCAGCTTTGGC	GGGCTGTGGA	AATCTCAAAA	TTTGAATGCA	TTTGAATGCA	TTTGTGCTAT
SEQ 15	CTCGAGCA - G	ATCCACGCGA	TCGTGAGGCG	CTGCGGAGCG	CTGCGGAGCG	TTGCCCTCAA	GG - - - CCGGC	TTTGAATGCA	TTTGAATGCA	TTTGTGCTAT
SEQ 17	GTTCGCGA - G	ATCAAGGATA	TCGTGCAAAA	GTGTCGGGTG	ACGGCGAGGA	TACGCGCGGA	GG - - - CCGGG	TTTGAATGCA	TTTGAATGCA	TTTGTGCTAT
SEQ 18	GTTCGCGA - G	ATCAAGGATA	TCGTGCAAAA	GTGTCGGGTG	ACGGCGAGGA	TACGCGCGGA	GG - - - CCGGG	TTTGAATGCA	TTTGAATGCA	TTTGTGCTAT
SEQ 20	GTGAGGA - G	ATTGAGGGAC	TCGTCAACAG	CTTTGTGGAC	GCTGCGAAGC	GTGCCATCGA	GG - - - CCGGC	GTGACATTTA	TTGAGATTCA	CGGCGCTCAC
SEQ 21	GTGAGGA - G	ATTGAGGGAC	TCGTCAACAG	CTTTGTGGAC	GCTGCGAAGC	GTGCCATCGA	GG - - - CCGGC	GTGACATTTA	TTGAGATTCA	CGGCGCTCAC
SEQ 23	GAAGAAGCA - A	TATGATGAAT	TAGTGAAGTA	GTGTTGTTGT	GCTGCGAAGC	GTGAGTTGTA	AA - - - TAGGT	TTTGAATGCA	TTTGAATGCA	TGGCGCTCAT
SEQ 25	ACCGAGGRRG	TCGAGGGTCT	GGGTGAAGAA	GTTCGCCGAG	TCGGCCAAAG	GGTCAAAATCG	A - - - GCTGGT	TTTGAATGCA	TTTGAATGCA	TGGCGCTCAT
SEQ 26	ACCGAGGRRG	TCGAGGGTCT	GGGTGAAGAA	GTTCGCCGAG	TCGGCCAAAG	GGTCAAAATCG	AG - - - CTGGT	TTTGAATGCA	TTTGAATGCA	TGGCGCTCAT
SEQ 28	AAGGAGGA - T	ATAGAGCAAC	TCAAGAGCGA	CTACGTGGAA	GGCGCAAAC	GAGCCATCCA	TG - - - CTGGT	TTTGAATGCA	TTTGAATGCA	TGGCGCTCAT
SEQ 29	AAGGAGGA - T	ATAGAGCAAC	TCAAGAGCGA	CTACGTGGAA	GGCGCAAAC	GAGCCATCCA	TG - - - CTGGT	TTTGAATGCA	TTTGAATGCA	TGGCGCTCAT
SEQ 32	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 34	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 36	AAGGAGGA - T	ATCGAGGAAC	TCAAGAATGA	CTTTCTGGCT	GCAGCAAAC	GAGCCAAWCCG	CGC - - - TGGT	TTTGAATGCA	TTTGAATGCA	TGGCGCTCAT
SEQ 37	AAGGAGGA - T	ATCGAGGAAC	TCAAGAATGA	CTTTCTGGCT	GCAGCAAAC	GAGCCAAWCCG	CGC - - - TGGT	TTTGAATGCA	TTTGAATGCA	TGGCGCTCAT
SEQ 39	TTGATGA - T	ATCGAGGCTT	TGAAGAAGGC	GTGTTGGAGAG	CGCGTCAAGC	GGGCAATTGAA	GGC - - - TGGG	TTTGAATGCA	TTTGAATGCA	TTTGTGCTAT
SEQ 41	CTCAGAGA - T	ATCAAGGAGA	TGGTCAAGA	CTGGGCGTGA	AGCGGCGTGA	AGC - - - GGGC	GTGGATGTTA	TTGAGATTCA	TTGAGATTCA	CGGCGGCTAC
SEQ 43	AAGGAGGA - T	ATTAAGGCGG	TGATTGAGGG	TTTGTGCCAC	ACGGCGGAGT	ACCTTGAAAA	GGC - - - CGGT	TTTGAATGCA	TTTGAATGCA	CGGCGGCTAC
SEQ 82	CTTGAAGA - G	ATTGAACATG	TGAAGGAGGA	TTTGTGTTCC	GAGCGGTGTA	GGGCGTTTGA	AG - - - CTGGG	TATGACTACG	TCGAAATCCA	CAGCGCTCAC
SEQ 84	ACCGAGGA - C	ATCAACAAGT	TGAAGAGCAA	ATTCTGTTAC	TCGGCACGAT	GGGCGTTTGA	AG - - - CTGGG	TATGACTACG	TCGAAATCCA	CAGCGCTCAC

	1201	1211	1221	1231	1241	1251	1261	1271	1281	1291
SEQ 1	GGATACCTGC	TGTGTCATT	CCTCTCGCG	GCGGCCAAC						
SEQ 2	GGATACCTGC	TGTGTCATT	CCTCTCGCG	GCGGCCAAC						
SEQ 4	GGCTATCTTC	TGATGTCGTT	CCTCTCCCT	GCGGTCAAC						
SEQ 5	GGCTATCTTC	TGATGTCGTT	CCTCTCCCT	GCGGTCAAC						
SEQ 7	GGCTATCTCA	TCAACGAGTT	CCTGAGCCG	GTACAGAA						
SEQ 9	GGGTATCTTC	TCTGTCCTT	CCTATCACC	TCTTCAAC						
SEQ 11	GGTTATTGGA	TAAATGAGTT	CTATAGTCT	ATTTCAAAC						
SEQ 13	GGATGTTTAA	TACACCAATT	TTTAAAGTAA	TTGACAAAC						
SEQ 15	GGCTACCTCA	TTTCCGAGTT	CTTGAGCCCC	ATCTCCAAC						
SEQ 17	GGATACCTGT	TGGCGCAGTT	CTTGAGCAAG	AAGACAAAC						
SEQ 18	GGATACCTGT	TGGCGCAGTT	CTTGAGCAAG	AAGACAAAC						
SEQ 20	GGTTACCTGA	TCACCGAGTT	CCTTTCGCC	CTATCAAAC	TAAGTGGAGA	TACTTTGTGT	GGGGCTGTGC	GCATACTCCC	TCGGGTGTGA	CTTCTATTAA
SEQ 21	GGTTACCTGA	TCACCGAGTT	CCTTTCGCC	CTATCAAAC						
SEQ 23	GGTTATCTTA	TATCGTCAAC	AGTTAGTCT	GCCACTAAT						
SEQ 25										
SEQ 26										
SEQ 28	GGATATCTAC	TGCATCAATT	CTTGAGTCCG	GTAAGCAAT						
SEQ 29	GGATATCTAC	TGCATCAATT	CTTGAGTCCG	GTAAGCAAT						
SEQ 32										
SEQ 34										
SEQ 36	GGATACKTGC	TTACACGATT	CTTGAGTCCA	GTCAAGTAAC						
SEQ 37	GGATACKTGC	TTACACGATT	CTTGAGTCCA	GTCAAGTAAC						
SEQ 39	GGATACCTCC	TCCACGAAAT	CATCTGCCTG	AGAGCAACA						
SEQ 41	GGGTACCTCA	TCCACGAAAT	CCTCTCACC	ATTACCAAC						
SEQ 43	GGTTACCTGC	TGGCCCAATT	CCTGTCGGA	ACAACCAAC						
SEQ 82	GGTTATCTTG	TTTCCAGCTT	CCTGTCGCC	GCCACCAAC						
SEQ 84	GGATACCTGA	TGCACTCGTT	CCTCAGCCG	TTGACCAAT						
	1301	1311	1321	1331	1341	1351	1361	1371	1381	1391
SEQ 1			AAACCGCAC	GGACCACTAC	GGCGGGTCTG	TCGAGAACC	CATCCGGCTG	TCTCTCGAGA	TTGCGCAGTT	GAATCGGGAC
SEQ 2			AAACCGCAC	GGACCACTAC	GGCGGGTCTG	TCGAGAACC	CATCCGGCTG	TCTCTCGAGA	TTGCGCAGTT	GAATCGGGAC
SEQ 4			ACGAGAAC	AGACGAGTAC	GGAGGCACTT	TTGAGAATCG	CATCCGGCTC	AGTCTGGAGA	TCGCCAAGCT	CACCCGCGAA
SEQ 5			ACGAGAAC	AGACGAGTAC	GGAGGCACTT	TTGAGAATCG	CATCCGGCTC	AGTCTGGAGA	TCGCCAAGCT	CACCCGCGAA
SEQ 7			AAGCGGAC	GGATGCGTAC	GGCGGGAGCT	TTGAGAACC	GACCCGGATC	GTGCGCGAGG	TTGCGCGGCG	TATTGCTGCG
SEQ 9			ACGCGCAC	CGACGAGTAC	GGCGGGTCTG	TTGAGAACC	CATCCGGCTC	TCTCTCGAAA	TCGCCAAGCT	CACCCGCGAA
SEQ 11			AAGAGAAC	AGATGAATAC	GGTGGCAGTT	TTGAAAATAG	AACCCAGATT	TTAAAGGAAG	TTATCGATAG	TGTTAAATCA
SEQ 13			AAGAGAGC	TGACCAATAC	GGGGGCTCAT	TTGAAAACAG	AGTTAGATTT	CTTTTACAAA	TAATTGAGAA	TATAAAACGA
SEQ 15			CAGCGTAC	CGACCACTAC	GGTGGCTCTT	TCGAGAACC	CACCCGGCTT	CTCCGCGAGA	TCATCTCGGC	CGTCCGCTCC
SEQ 17			AGGCGCGG	GGATGAGTAT	GGCGGGTCTG	TCGAGAACC	GCGCGAGGAT	GTGCGGGAGA	TTATTAAGGA	GTGCGAGGAG
SEQ 18			AGGCGCGG	GGATGAGTAT	GGCGGGTCTG	TCGAGAACC	GCGCGAGGAT	GTGCGGGAGA	TTATTAAGGA	GTGCGAGGAG
SEQ 20	CATTTTATTT	CCTGGCACCG	AGAAACGGAC	AGACAAGTAC	GGCGGCACTT	TTGAGAACC	CACCCGGGTC	CTGATCGATA	TTATCAAGGC	CGTCCGGGCA
SEQ 21			AAACGGAC	AGACAAGTAC	GGCGGCACTT	TTGAGAACC	CACCCGGGTC	CTGATCGATA	TTATCAAGGC	CGTCCGGGCA
SEQ 23			GACCGCAA	TGACAAGTAT	GGTGGGACAT	TTGAGAAAC	TATTTTGTTT	CCTATGGAAG	TTGTCTCAAT	TGTTCTGTAA
SEQ 25										
SEQ 26										
SEQ 28			CAAAGAAC	CGACGAGTAT	GG					
SEQ 29			CAAAGAAC	CGACGAGTAT	GG					
SEQ 32			AAC	CGACGAGTAT	GGTGGCAGTT	TCGAGAACC	TATCAGAGTT	GTCTTGGAAA	TCCCTGACCT	CATCCGCGCT
SEQ 34										
SEQ 36			CAAAGAAC	CGATGAGTAT	GGTGGCAGCT	TCGAGAACC	TATCAGAGTT	GTCTTGGAGA	TCATTG	
SEQ 37			CAAAGAAC	CGATGAGTAT	GGTGGCAGCT	TCGAGAACC	TATCAGAGTT	GTCTTGGAGA	TCATTG	
SEQ 39			CCAGGACC	GACAAGTACG	GGGAAACCG	CACCTCGTCTG	ACAATGGAAA	GTGCTCGACC		TTGTCCGCG
SEQ 41			CGCCGGAC	AGATTCTTAC	GGCGGTTCTT	TCGAAAACCG	TACCCGTCTA	CTCATTTGAAA	TCGTAACAGC	CGTCCGAGCC
SEQ 43			CAGCGCAC	CGACGAGTAC	GGCGGCACTT	TCGAAAACCG	CATCCGGGCTA	ATCTCTGAGG	TCACGGCGCA	GGTCCGCGAG
SEQ 82			AAGCGTAC	CGACAAGTAC	GGAGGTAGCT	TCGAGAACC	AGTGGCGCTT	GCTCTCGAGA	TTGTCGAGGC	TGCAAGAGCT
SEQ 84			CAGCGTAC	CGACGAGTAC	GGCGGTAGCC	TGGAGAACCG	CGCTCGATT	CTGCTCAACG	TTGCCGCTCG	AATCCGCGCA
	1401	1411	1421	1431	1441	1451	1461	1471	1481	1491
SEQ 1	GCGGTGCGCC	CTCATGTGCC			GTCTTT	CCTGCGCAIT	TCGGGCTCGG	ACTGGTGCGA	GGAGACCTTG	CCGGA
SEQ 2	GCGGTGCGCC	CTCATGTGCC			GTCTTT	CCTGCGCAIT	TCGGGCTCGG	ACTGGTGCGA	GGAGACCTTG	CCGGA
SEQ 4	AATGTGCCCA	AGGATATGCC			GTCTTT	CCTGCGGCTC	TCCGCCACCG	ATTGGCTGGA	GGAGGTGCGA	CCGAA
SEQ 5	AATGTGCCCA	AGGATATGCC			GTCTTT	CCTGCGGCTC	TCCGCCACCG	ATTGGCTGGA	GGAGGTGCGA	CCGAA
SEQ 7	GTGATTCCCG	AGGGGATGCC			CTGTTT	TCTGCGTATC	AGCGCCACCG	AGTGGTTGGA	GGGTGAGCGG	GTGGC
SEQ 9	GCGGTGCGCC	CCAACGTGCC			GTCTTT	TCTGCGTATC	TCCGCCACCG	ACTGGATGGA	GGAGACCTTG	CCGGA
SEQ 11	AGTATTCCAA	ACGATGTGCC			GTCTTT	TCTGAGAAAT	TCTGCTGCTG	AAAATAGTCC	TGATCCA	
SEQ 13	AAGATAGAAA	CA	CC		ATTTT	CTTAAAGTTT	CCAATGTGAG	ATAATTTGAG	TGATCCG	
SEQ 15	GTATCTCCCG	AGGACATGCC			CTCTT	CGTCCGTGTC	TCCGCCACCG	AGTGGATGGA	GTACACC	
SEQ 17	CAGGTGACTG	AGGCGGTGGG	TGAAGAGGAG	GCGAAGAAAT	TTGTGTTGGG	AATCAAGCTG	AACAGTGCAG	ATTGGCAGGC	GGGACGCGAT	GGA
SEQ 18	CAGGTGACTG	AGGCGGTGGG	TGAAGAGGAG	GCGAAGAAAT	TTGTGTTGGG	AATCAAGCTG	AACAGTGCAG	ATTGGCAGGC	GGGACGCGAT	GGAAAG
SEQ 20	GTGATTCCCG	AGGAGATGCC			CTCTT	CGTCCGAATC	TCCGCCACCG	AATGGATGGA	GTACCCGCGC	
SEQ 21	GTGATTCCCG	AGGAGATGCC			CTCTT	CGTCCGAATC	TCCGCCACCG	AATGGATGGA	GTACCCGCGC	
SEQ 23	GCAATTCCAG	ATAGTATGCC			TTGTT	TTATAGAGTA	ACGGCTACAG	ATTGGTTGCC	CAAAGGACAA	
SEQ 25										
SEQ 26										
SEQ 28										
SEQ 29										
SEQ 32	GCCATCCCCG	AAACTACACC	T		GTCTT	CGTTCGTGTC	AGTGCAACTG	ATTGGTTGGA	GTTTGACTCT	CAATTCAAAG
SEQ 34										
SEQ 36										
SEQ 37										
SEQ 39	CATT									
SEQ 41	GCGATGCCCT	CCAGCATGCC	T		CTCTT	CCTCCGCTC	TCTCTACAG	AATGGATGGA	AGATACCGAC	ATCGGC
SEQ 43	CGGACGAGCA	AGAATTTTAT	C		CTCGG	CATCAAAAT	AACAGCGTCG	AGTTTCCAGGA	GAAG	
SEQ 82	GTATGCTCTG	AGGACATGCC			TTGTT	CACCTGCATC	AGTGGAACTG	ACTGGCTGGA	GAACAAACCT	GAG
SEQ 84	GAATTCCTCA	ACAAGGGT			CTCTG	GGTGGCGCTC	AGTCCACCG	ACTGGGCGGA	CCAAGCGCAC	CAA

	1801	1811	1821	1831	1841	1851	1861	1871	1881	1891
	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 1	---AAGCT	GCTGGTTGCC	GCCGTGGGTG	CCATCACC					AACG	GCAAGCAGGC
SEQ 2	---AAGCT	GCTGGTTGCC	GCCGTGGGTG	CCATCACC					AACG	GCAAGCAGGC
SEQ 4	---AAACT	CGCAGTGGCA	TCAGTGGGTA	TGATTGCC					AGCG	CGCATTTCGC
SEQ 5	---AAACT	CGCAGTGGCA	TCAGTGGGTA	TGATTGCC					AGCG	CGCATTTCGC
SEQ 7	---GCTGG	CGCGTCGACT	CTTGTTGGGT	CTGTAGGTCT	GATCACCAGT	TCGGAACAGG	CGAGGGGACT	AGTTCAGGGA	GCGGACGAG	CGACTGCAGC
SEQ 9	---AAGCT	CCTTGTTCGC	ACGGTGGGCA	CGATCAGC					AACG	GTAAGCAGGC
SEQ 11	---AAGTT	ATTGGTCACT	TGCGTTGGGT	GGCTTGAA					A	AAGATCCTGA
SEQ 13	---CGATG	TTTGATTCGCA	TGCAGTGGAG	GATTAGAT					C	GAGACATATT
SEQ 15	GCAAGCAGCT	CTCTGTCGGT	GCGCTCGGCT	TGGTCACC	TCG	GCTGAGATCG	CCAAGGAGAC	CGTCCAGGAG	AAGGAGGATG	GCAGAGTCACT
SEQ 17	TCCCCAAGCT	TCCTCTCATG	GTCACCGCGC	GCTTCCGC					ACTC	GTACAGGCACT
SEQ 18	TCCCCAAGCT	TCCTCTCATG	GTCACCGCGC	GCTTCCGC					ACTC	GTACAGGCACT
SEQ 20	---AGGTT	GCTCATAGGC	GCGGTCGGCA	ACATCAAC					ACGG	CTGACATTGC
SEQ 21	---AGGTT	GCTCATAGGC	GCGGTCGGCA	ACATCAAC					ACGG	CTGACATTGC
SEQ 23	---AT	ACTACTTGGC	GCTGTCGGAA	TGATCAGG					GATG	GTCCTTACGGC
SEQ 25										
SEQ 26										
SEQ 28										
SEQ 29										
SEQ 32	---AGTGT	ACTTGTTTCA	GCAGTAGGTG	GAATCAAG					A	CTGGACATCT
SEQ 34										
SEQ 36										
SEQ 37										
SEQ 39										
SEQ 41										
SEQ 43	---ATGGT	GGTCTACACC	ACCGCGCGCT	TCAAGAGC					GTGGGGC	CCATGGTTCGA
SEQ 82	---AAGAT	GTTGATCAGC	ACTGTTGGTA	GCATCAAG					ATAG	GTACCCTTGC
SEQ 84	--ATCGAACC	CGACGCGTCC	AAACGCATGC	TCGTCGGGG					CCGTGG	GAATGATGGA
	1901	1911	1921	1931	1941	1951	1961	1971	1981	1991
	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 1	GAATCAG	---ATTCTAG	AGGAGCAG							
SEQ 2	GAATCAG	---ATTCTAG	AGGAGCAG							
SEQ 4	CAATTCC	---TTGTTGG	AGAAGGAC							
SEQ 5	CAATTCC	---TTGTTGG	AGAAGGAC							
SEQ 7	CGAGGCCAATG	CTGTGGGAC	CTGAACCC							
SEQ 9	GAACAAG	---CTGCTTG	AGGAGGAG							
SEQ 11	ATTGCTCAAC	AAATATTTAG	AAGAAGGA							
SEQ 13	TAAACTCGAT	GAGTTATTATG	CTAATGGT							
SEQ 15	CATCCAGCGC	GAGAACGGCG	CCAAGACT							
SEQ 17	GGAGGCC	---GCTTTGG	AATCCGAT							
SEQ 18	GGAGGCC	---GCTTTGG	AATCCGAT							
SEQ 20	GCGCGATGTC	GTGGATGAGC	AGGGCGCCGA	GAAGGTGGCC	GAGGCCAAGC	AGACGCATGA	CACCATCGAG	GTCTGTGAGC	AATCATATGG	CGGCAAGACC
SEQ 21	GCGCGATGTC	GTGGATGAGC	AGGGCGCCGA	GAAGGTGGCC	GAGGCCAAGC	AGACGCATGA	CACCATCGAG	GTCTGTGAGC	AATCATATGG	CGGCAAGACC
SEQ 23	GAATGAAATC	CTAGAAAGTG	AAAAGCT							
SEQ 25										
SEQ 26										
SEQ 28										
SEQ 29										
SEQ 32	TGCTGAA	---GAGGTTT	TGCAATCT							
SEQ 34										
SEQ 36										
SEQ 37										
SEQ 39										
SEQ 41										
SEQ 43	GCGCGTCGAC	GGCGTCGATG	GG							
SEQ 82	GGAGGAG	---ATCATCG	CTGAGGAGA	GGACGATACC						
SEQ 84	AGGTTCC	---TACGATT	CGCCCAAC							
	2001	2011	2021	2031	2041	2051	2061	2071	2081	2091
	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 1	GATATCGAGC	TTGCGCTGGT	TGGCCGTGGG	TTCCAGAAGG	ATCCCGGTCT	GGCCTGGAGC	TTTGCTCAGC	ACCTCGGCGT	C	
SEQ 2	GATATCGAGC	TTGCGCTGGT	TGGCCGTGGG	TTCCAGAAGG	ATCCCGGTCT	GGCCTGGAGC	TTTGCTCAGC	ACCTCGGCGT	C	
SEQ 4	GGACTGGACC	TTGTGCTGGT	TGGAAGTGGC	TTCCAGAAGA	ACCCGGGGCT	GGTGTGGGCG	TGGGCCGACG	AGCTGAATGT	A	
SEQ 5	GGACTGGACC	TTGTGCTGGT	TGGAAGTGGC	TTCCAGAAGA	ACCCGGGGCT	GGTGTGGGCG	TGGGCCGACG	AGCTGAATGT	A	
SEQ 7	AAGCGGGATG	CCATTCTGAT	AGCCCGTCAG	TTCTTGCGCG	AGCCAGAATG	GGTGTTTTC	ACGGCGAGAA	AGTTGGGCGT	G	
SEQ 9	GGATTGGATG	TTGCGCTTGT	GGGACGTGGT	TTCCAGAAGG	ATCCCGGTCT	GGCGTGGACT	TTCCGCGCAGC	ATCTTGATGT	T	
SEQ 11	ACATTTGATC	TGTCTTTGAT	CGGTAGAGGA	TTTTTAAGAA	ATCCAGGTTT	GGTATGGGAG	TTTGCCGATA	AACCTTGGTG	T	
SEQ 13	GACTTTGATA	TAGCATTTGAT	AGGTAAGGA	TTTCTCAAAA	ACACTGGATT	GATCAGCCGT	ATTGCTGACC	AATTGCAAGC	A	
SEQ 15	CGTGCCGATA	TGGTCTTGT	TGCCAGGCAG	TTCTTTAAGG	AGCCCGAGTT	CGTCTCTACT	GTGCGCGACG	AGTTGGGTGT	T	
SEQ 17	GATTTCGCACA	TGATCGGTAT	CGGACGCCCG	GCCATCATCA	ACCCCTTCGT	TCCCGCCAA	TTGATCCTCA	ACCCGGAGGT	G	
SEQ 18	GATTTCGCACA	TGATCGGTAT	CGGACGCCCG	GCCATCATCA	ACCCCTTCGT	TCCCGCCAA	TTGATCCTCA	ACCCGGAGGT	G	
SEQ 20	AAGGCGGATC	TGTCCTTCAT	TGCTGCCAG	TTCTTGCGCG	AGCCTGAGTT	TGTGCTGAGG	ACGGCGCATA	ACCTTGGGGT	C	
SEQ 21	AAGGCGGATC	TGTCCTTCAT	TGCTGCCAG	TTCTTGCGCG	AGCCTGAGTT	TGTGCTGAGG	ACGGCGCATA	ACCTTGGGGT	C	
SEQ 23	-----GATG	TTACTTTTGT	CGCAAGGGAG	TTCTTAAGGA	ACCCGTCGTT	GGTGCTAGAC	AGCGCAAGC	AGTTGGGTGA	A	
SEQ 25										
SEQ 26										
SEQ 28										
SEQ 29										
SEQ 32	GGTATCGACA	TTGTGAGGGC	TGGAAGTTGG	TTCCAACAGA	ATCCTGGTCT	GGTTCGAGCT	TTTGCTAAGC	AGCTTGGCGT	G	
SEQ 34										
SEQ 36										
SEQ 37										
SEQ 39										
SEQ 41										
SEQ 43	-----	--ATAGGCAT	CGGGCGCGCA	GCCGGTTCGG	AGCCCGGACCT	CGCCAAGGAC	ATCATCGCGG	GCAAGGTGTC	CAGCATTATC	AAATACGCCA
SEQ 82	CCCTTGGATC	TTGTGGCTTC	AGGCCGTCG	TTCCAGAAGA	ACACTGGACT	TGTTTGCTCA	TGGGCTGACG	ATCTGAACAC	T	
SEQ 84	GGCCAACAGC	CGAGCCAGAT	TGGCAAGTGT	GCCGAGCAGT	CGATTCCAGAG	CGGAGAGTGT	GATGCGGTAC	TGTTGGCAGC	T	

	2401	2411	2421	2431	2441	2451	2461	2471	2481	2491
SEQ 1	CGTCCTCTTA	AGTTTCTCCG	TCATTCTGTC	TATTCTACTC	CAATCGCAAC	GCAATGGCGAC	CACGGATCGA	GTCGAATTTC	TCCGTCGTTT	GTATCTGATC
SEQ 2	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 4	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 5	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 7	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 9	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 11	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 13	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 15	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 17	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 18	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 20	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 21	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 23	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 25	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 26	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 28	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 29	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 32	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 34	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 36	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 37	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 39	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 41	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 43	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 82	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 84	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
	2501	2511	2521	2531	2541	2551	2561	2571	2581	2591
SEQ 1	AATATAAAAA	GCGGGGAATG	GCTTGACCCC	GCGCAGAATG	TCGATCTCTT	CGCAAACCTCT	CGGTGTATAG	GACGCTCAGC	AACGATCAAG	G
SEQ 2	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 4	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 5	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 7	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 9	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 11	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 13	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 15	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 17	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 18	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 20	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 21	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 23	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 25	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 26	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 28	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 29	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 32	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 34	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 36	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 37	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 39	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 41	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 43	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 82	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
SEQ 84	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Figure 2. A multiple alignments of the 2031 OR nucleic acid sequence from *A. fumigatus* (SEQ 1,2) along with related 2031 ORs from other fungi and bacteria (see also Example 4). Regions 1-11, marked with * or #, refer to regions conserved at the amino acid level between Ors but not OYEs.

Fungal 2031 ORs are given by SEQ ID No.: SEQ ID Nos. 1, 2, 4, 5, and 7, *A. fumigatus*; SEQ ID No. 9, *A. nidulans*; SEQ ID Nos. 11 and 13, *C. albicans*; SEQ ID Nos. 15, 17 and 18, *N. crassa*; SEQ ID Nos. 20, 21 and 43, *M. grisea*; SEQ ID No. 23 (NP_595868), *S. pombe*; SEQ ID Nos. 25 and 26, *C. trifolii*; SEQ ID Nos. 28, 29, 31, 32 and 34, *F. sporotrichioides*; SEQ ID Nos. 36, 37 and 82, *F. graminearum*; SEQ ID Nos. 39 and 41, *M. graminicola*; SEQ ID No. 84, *U. maydis*.

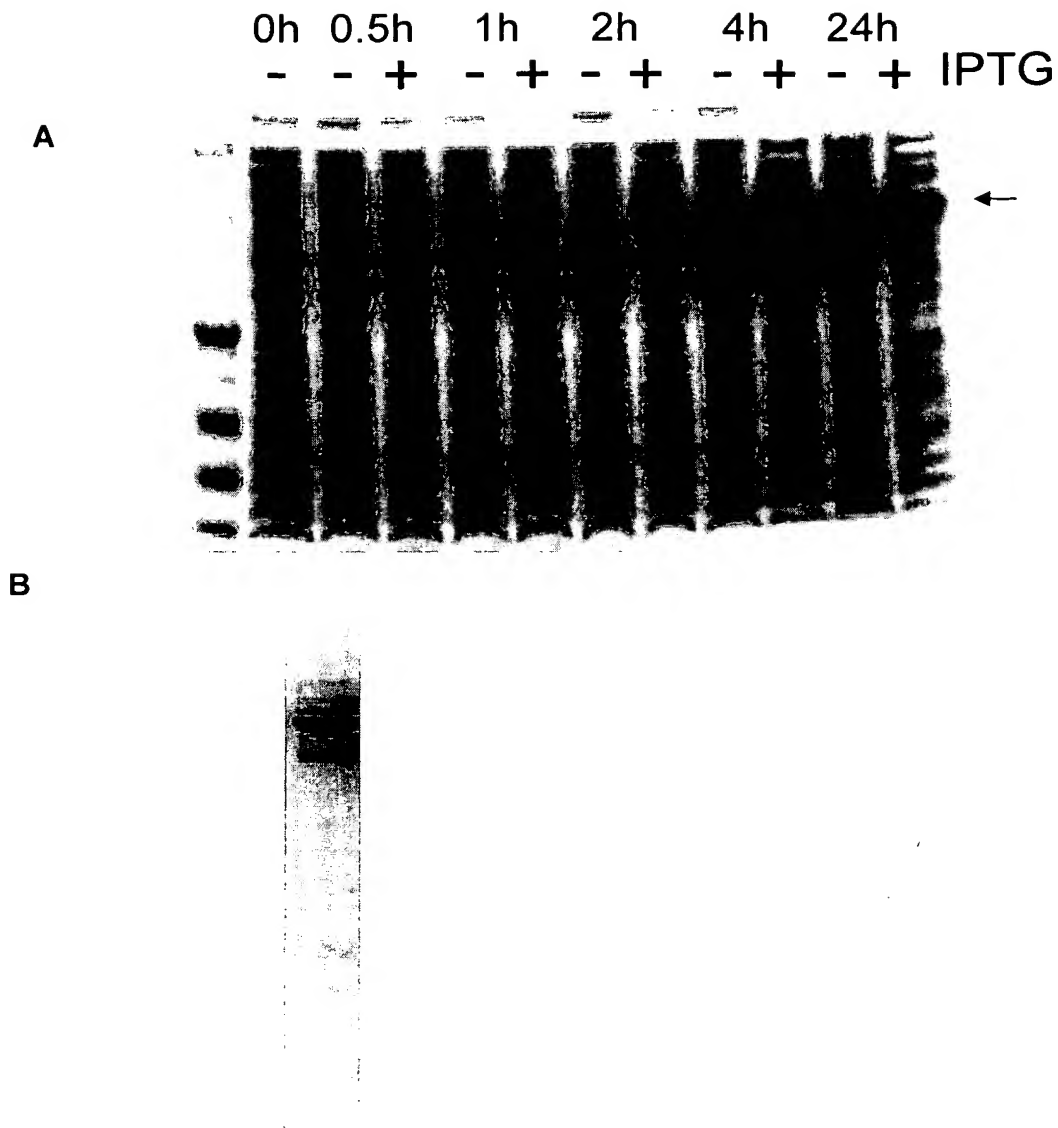


Figure 3. Recombinant 2031 OR. (A) Time course of recombinant 2031 OR induction over 24 hours after the addition of IPTG (samples without IPTG are also shown). The gel was stained with coomassie; A prominent band of the correct molecular weight (marked with an arrow) is seen. (B) Coomassie stained gel showing purified recombinant 2031.

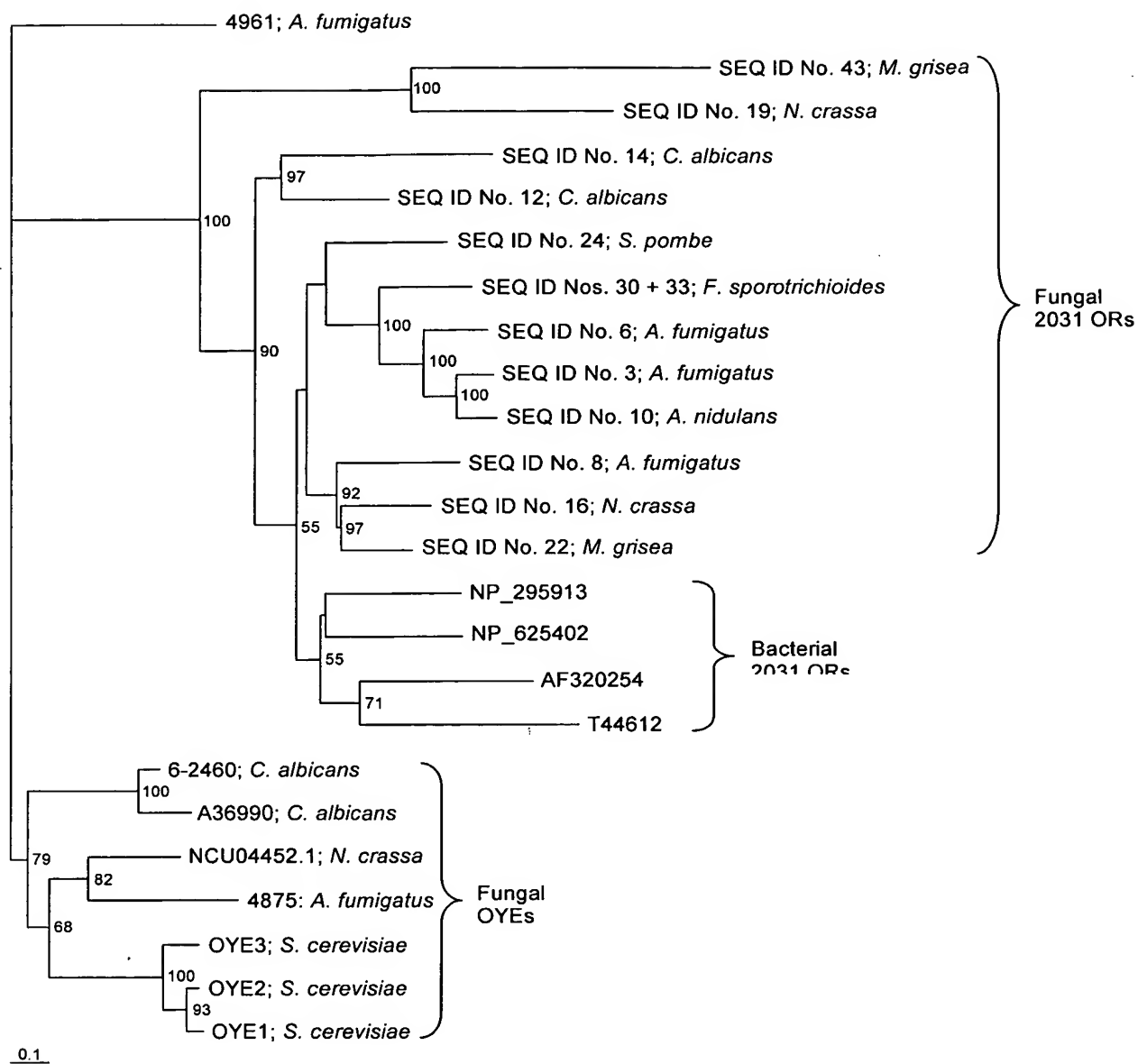


Figure 4. Phylogenetic tree showing relationships between *A. fumigatus* 2031 OR and similar proteins. This demonstrates a 2031 OR clade, which can be distinguished from the OYE proteins.

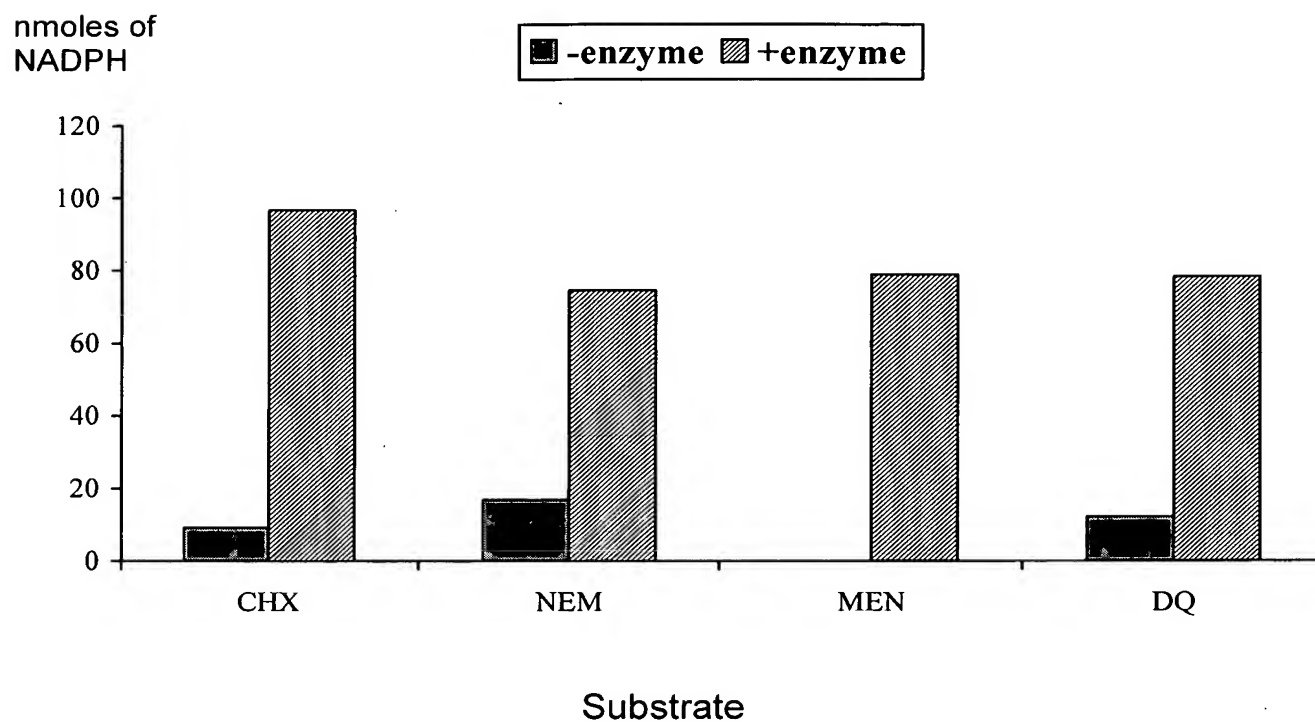


Figure 5: NADPH dehydrogenase activity of recombinant 2031 OR with cyclohexenone (CHX), N-ethylmaleimide (NEM), menadione (MEN) or duroquinone (DQ) as substrates. Final concentrations in the assay were as follows: 500 μ M substrate, 120 μ M NADPH, 1 μ g/200 μ L 2031 OR.

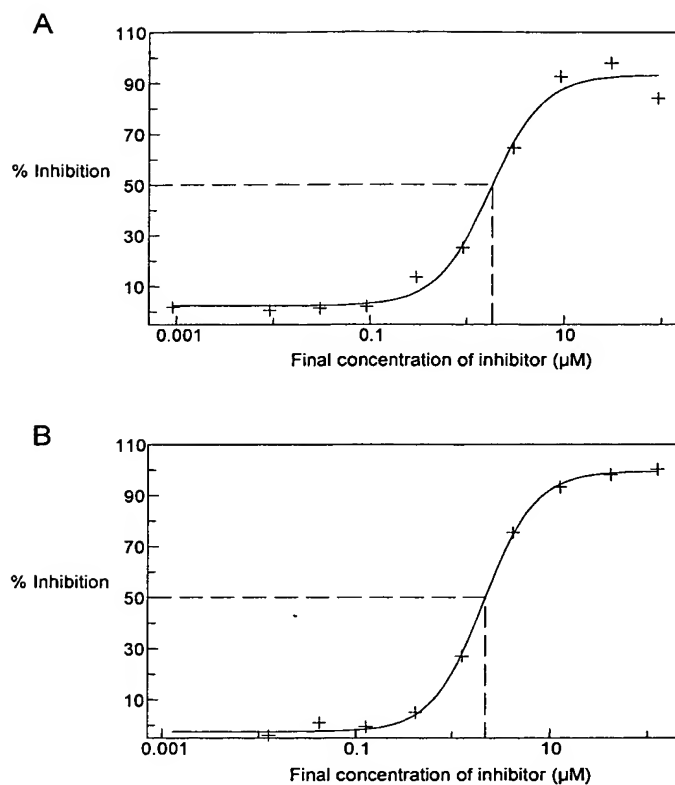


Figure 6: Inhibition of 2031 OR function by two inhibitors (shown in **A** and **B**) identified by high-throughput screening.